Journal of Multivariate Analysis 101 (2010) 223–239

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Testing quasi-independence for truncation data

Takeshi Emura, Weijing Wang *

National Chiao-Tung University, Hsin-Chu, Taiwan, ROC

ARTICLE INFO

Article history:
Received 27 August 2008
Available online 3 August 2009

AMS subject classifications:
62N01
62N03

Keywords:
Conditional likelihood
Kendall's tau
Mantel–Haenszel test
Power
Right-censoring
Survival data
Two-by-two table

ABSTRACT

Quasi-independence is a common assumption for analyzing truncated data. To verify this condition, we propose a class of weighted log-rank type statistics that include existing tests proposed by Tsai (1990) and Martin and Betensky (2005) as special cases. To choose an appropriate weight function that may lead to a more power test, we derive a score test when the dependence structure under the alternative hypothesis is modeled via the odds ratio function proposed by Chaieb, Rivest and Abdous (2006). Asymptotic properties of the proposed tests are established based on the functional delta method which can handle more general situations than results based on rank-statistics or U-statistics. Extension of the proposed methodology under two different censoring settings is also discussed. Simulations are performed to examine finite-sample performances of the proposed method and its competitors. Two datasets are analyzed for illustrative purposes.

1. Introduction

Truncated data are commonly seen in studies of biomedicine, epidemiology, astronomy and econometrics. Such data occur when the variables of interest can be observed if their values satisfy certain criteria. In this article, we discuss the situation that a pair of lifetime variables (X, Y) can be included in the sample only if $X \le Y$. The variable Y is said to be left-truncated by X and X is right-truncated by Y. Sometimes, external censoring also happens due to subjects' withdrawal or the end-of-study effect. Here we allow that Y is subject to right-censoring by another variable C. Hence one observes X, $Z = Y \wedge C$ and $\delta = I(Y \le C)$ subject to $X \le Z$, where $a \wedge b = \min(a, b)$ and $I(\cdot)$ is the indicator function. Left-truncated and right-censored data consist of $\{(X_i, Z_i, \delta_i)(i = 1, \ldots, n)\}$, replications of (X, Z, δ).

Truncation often occurs when a subject can be recruited according to a certain sampling criterion [1]. For example in the study of transfusion-related AIDS discussed in Lagakos, et al. [2], infected people could be included in the sample only if they developed AIDS within the study period. Accordingly the incubation time X was subject to right-truncation by the lapse time Y measured from infection to the recruitment time. In this design, a subject with the incubation time exceeding the lapse time $(X > Y)$ would never be observed. Another example is the survival analysis for residents in the Channing House retirement community in Palo Alto, California [3–5]. This sample cannot represent the general population since only those who had lived long enough to enter the retirement center could be observed. Hence the lifetime Y was left-truncated by the entry age X. Notice that a truncated subject with $X > Y$ is completely missing and even its existence is unknown.

Any statistical analysis for data subject to truncation requires making some assumption about the association between X and Y. Independence between X and Y is the most common assumption [3,5,6,2,7,1,8]. This assumption has been relaxed by Tsai [9] to a weaker condition of quasi-independence which can be formulated as follows:

$$H_0 : \pi(x, y) = F_X(x)S_Y(y)/c_0 \quad (x \le y), \tag{1}$$

* Corresponding author. Fax: +86 3 5728745.
E-mail addresses: temura@stat.nctu.edu.tw (T. Emura), wjwang@stat.nctu.edu.tw (W. Wang).

0047-259X/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2009.07.006

Table 1
2×2 table for truncated data without censoring.

	$Y = y$	$Y > y$	
$X = x$	$N_{11}(\mathrm{d}x, \mathrm{d}y)$		$N_{1\bullet}(\mathrm{d}x, y)$
$X < x$			
	$N_{\bullet 1}(x, \mathrm{d}y)$		$R(x, y)$

where $\pi(x, y) = \Pr(X \le x, Y > y | X \le Y)$ and F_X and S_Y are arbitrary right-continuous distribution and survival functions respectively, and c_0 is the constant satisfying $c_0 = -\iint_{x \le y} \mathrm{d}F_X(x)\mathrm{d}S_Y(y)$. The joint function $\pi(x, y)$ is defined in the upper wedge ($x \le y$) and, under H_0, it can be factorized into the product of two marginal functions F_X and S_Y. Since the behavior of these functions in the lower wedge ($x > y$) is not specified, F_X and S_Y may not be equal to the true distribution and survival functions of X and Y respectively [10]. The assumption of quasi-independence in (1) is weaker than independence. Thus rejection of H_0 implies rejection of independence between X and Y but not vice versa. Many nonparametric methods for truncation data are still valid under H_0. If X and Y are truly independent, H_0 holds and then $F_X(x) = \Pr(X \le x)$, $S_Y(y) = \Pr(Y > y)$ and $c_0 = \Pr(X \le Y)$.

Unlike independent censorship which cannot be verified, quasi-independence is a testable assumption [9]. Tsai [9] proposed the first test on H_0 by defining a conditional version of Kendall's tau and then using its empirical estimator as the test statistics. Martin and Betensky [11] extended Tsai's idea to more complicated truncation structures in which the properties of U-statistics are applied in variance estimation and large-sample analysis. Chen, Tsai and Chao [12] constructed their test based on a conditional version of Pearson correlation coefficient.

In this article, we propose different methods for testing H_0. Specifically based on a series of 2×2 tables suitable for describing truncated data, we construct weighted log-rank type tests. We also show that the tests of Tsai [9] and Martin and Betensky [11] can be viewed as our special cases with different forms of weight. To choose a good weight that leads to a more powerful test, we propose a score test that utilizes some distributional properties of the 2×2 tables. In particular, the odds ratio function proposed by Chaieb, Rivest and Abdous [10] is adopted to model the dependence structure under the alternative hypothesis. The existing testing procedures also differ in the way of estimating the variance of the corresponding test statistic. Here we adopt the functional delta method which can handle flexible weight functions and hence is a more powerful tool than the techniques based on rank-statistics or U-statistics.

The paper is organized as follows. In Section 2, we propose the main methodology by temporarily ignoring censoring. In Section 3, we derive the score test and suggest a model selection method. Large-sample properties are examined in Section 4. Modifications of all the results to account for the presence of right-censoring are presented in Section 5. Section 6 contains numerical analysis including data analysis and simulation studies. Concluding remarks are given in Section 7.

2. The proposed method without censoring

To illustrate the main idea, we temporarily ignore right-censoring by letting $C = \infty$. Observed data can be expressed as $\{(X_j, Y_j) : (j = 1, \dots, n)\}$ subject to $X_j \le Y_j$.

2.1. Constructing the test statistics based on two-by-two tables

Adapt to the nature of truncation, we can construct the following 2×2 table at an observed failure point (x, y) for $x \le y$. The cell counts and marginal counts in Table 1 are defined as

$$N_{11}(\mathrm{d}x, \mathrm{d}y) = \sum_j I(X_j = x, Y_j = y), \qquad N_{\bullet 1}(x, \mathrm{d}y) = \sum_j I(X_j \le x, Y_j = y),$$
$$N_{1\bullet}(\mathrm{d}x, y) = \sum_j I(X_j = x, Y_j \ge y), \qquad R(x, y) = \sum_j I(X_j \le x, Y_j \ge y).$$

Under H_0 and conditional on the marginal counts, the cell count $N_{11}(\mathrm{d}x, \mathrm{d}y)$ follows the hyper-geometric distribution with

$$E(N_{11}(\mathrm{d}x, \mathrm{d}y) | N_{1\bullet}, N_{\bullet 1}, R) = \frac{N_{1\bullet}(\mathrm{d}x, y) N_{\bullet 1}(x, \mathrm{d}y)}{R(x, y)}. \tag{2}$$

To test quasi-independence, we propose the following weighted log-rank type statistics:

$$L_W = \iint_{x \le y} W(x, y) \left\{ N_{11}(\mathrm{d}x, \mathrm{d}y) - \frac{N_{1\bullet}(\mathrm{d}x, y) N_{\bullet 1}(x, \mathrm{d}y)}{R(x, y)} \right\}, \tag{3}$$

where $W(x, y)$ is a weight function. Motivated by the G^ρ class discussed in Harrington and Fleming [13,14], we consider a sub-class of L_W with a particular form of $W(x, y)$ which can be written as

$$L_\rho = \iint_{x \le y} \hat{\pi}(x, y-)^\rho \left\{ N_{11}(\mathrm{d}x, \mathrm{d}y) - \frac{N_{1\bullet}(\mathrm{d}x, y) N_{\bullet 1}(x, \mathrm{d}y)}{R(x, y)} \right\}, \tag{4}$$

where $\hat{\pi}(x, y-) = R(x, y)/n$ and $\rho \ge 0$ is a pre-specified constant.

The statistics L_W is nonparametric in the sense that no distributional assumption about the joint distribution of (X, Y) is made. However such information would be helpful for choosing an appropriate weight or the value of ρ in (4) which may lead to a more powerful test. In Section 3, we derive a score test that utilizes the information of the underlying association structure provided by the 2×2 tables.

2.2. Relationship with existing tests

The tests proposed by Tsai [9] and Martin and Betensky [11] are both related to a conditional version of Kendall's tau defined as

$$\tau_a = E\{\mathrm{sgn}(X_i - X_j)(Y_i - Y_j)|A_{ij}\},$$

where $\mathrm{sgn}(x)$ is defined to be $-1, 0$, or 1 if $x < 0, x = 0$, or $x > 0$ respectively, $A_{ij} = I\{\check{X}_{ij} \le \tilde{Y}_{ij}\}, \check{X}_{ij} = X_i \vee X_j$ and $\tilde{Y}_{ij} = Y_i \wedge Y_j$. Note that when the event A_{ij} occurs, $(\check{X}_{ij}, \tilde{Y}_{ij})$ is located in the observable region $\{(x, y) : 0 < x \le y < \infty\}$ and hence τ_a is well defined under the truncation setting. Under quasi-independence, Tsai [9] showed that $\tau_a = 0$.

An empirical estimator of τ_a can be used for testing H_0. Specifically Tsai [9] and Martin and Betensky [11] both considered the statistics

$$K_a = \sum_{i<j} I\{A_{ij}\}\mathrm{sgn}\{(X_i - X_j)(Y_i - Y_j)\} \tag{5}$$

but proposed different ways of calculating the variance of K_a. For example in the absence of ties, by writing K_a as the sum of conditionally independent rank variables, Tsai [9] was able to utilize rank-based results to derive the conditional variance of K_a explicitly. Martin and Betensky [11] recognize the fact that K_a is a U-statistic and then derive a more general variance formula which can handle tied data. The statistic K_a has been extended to account for censoring [9] or even more complicated data structures [11].

Now we compare the proposed test statistics L_W in (3) with K_a in (5). To simplify the analysis, assume that the data have no ties so that the values of $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ are all distinct. In such a case $N_{\bullet 1}(x, dy) = N_{1\bullet}(x, dy) = 1$ for all tables of interest and the expected value in (2) becomes $1/R(x, y)$. It can be shown that

$$L_W = -\sum_{i<j} I\{A_{ij}\} \frac{W(\check{X}_{ij}, \tilde{Y}_{ij})}{R(\check{X}_{ij}, \tilde{Y}_{ij})} \mathrm{sgn}\{(X_i - X_j)(Y_i - Y_j)\}. \tag{6}$$

The proof of the above equation is given in Appendix C under a more general setting that includes right-censoring. By setting $W(x, y) = R(x, y)/n$, we get

$$L_{\rho=1} = \iint_{x \le y} \frac{R(x, y)}{n} \left\{N_{11}(dx, dy) - \frac{N_{1\bullet}(dx, y)N_{\bullet 1}(x, dy)}{R(x, y)}\right\} = -\frac{K_a}{n}. \tag{7}$$

Eq. (6) implies that L_W is also a U-statistic if $W(x, y)/R(x, y)$ is a deterministic function. However if we prefer a flexible weight function that may lead to a more powerful test, the technique of U-statistics is no longer applicable for variance estimation and large-sample analysis. Accordingly in Section 4, we will use the functional delta method to establish asymptotic properties of L_W.

3. Conditional score test

3.1. Construction of conditional likelihood

As mentioned above, the weight function in (3) affects the power of L_W which depends on the dependence structure under the alternative hypothesis. The Clayton model [15], characterized by the constant odds ratio function [16,17], is perhaps the most popular choice for describing bivariate lifetime variables. The class of Archimedean copula (AC) models, which include the Clayton model and the bivariate frailty family [17] as special cases, provide a systematic framework to describe the dependence for multivariate random variables [18]. These concepts are modified by Chaieb, et al. [10] in analysis of truncated data. Here we also adopt their proposal.

We assume that $\pi(x, y) = \Pr(X \le x, Y > y|X \le Y)$ is differentiable and hence the data have no ties. Chaieb, et al. [10] modified the odds ratio function suitable for truncated data as follows:

$$\vartheta(x, y) = \frac{\pi(x, y) \cdot \partial^2\pi(x, y)/\partial x\partial y}{\partial\pi(x, y)/\partial x \cdot \partial\pi(x, y)/\partial y}.$$

Under quasi-independence, $\vartheta(x, y) = 1$ for all $0 < x \le y$. It should be noted that the case of $\vartheta(x, y) < 1$ implies positive association while $\vartheta(x, y) > 1$ implies negative association between the two truncated variables.

The information of $\vartheta(x, y)$ is contained in the summary statistics of Table 1. Given the marginal counts, $N_{11}(dx, dy)$ follows a Bernoulli distribution with

$$\Pr(N_{11}(dx, dy) = 1 | N_{1\bullet} = N_{\bullet 1} = 1, R = r) = \frac{\vartheta(x, y)}{r - 1 + \vartheta(x, y)}.$$

This distributional result can be further utilized to construct a score test under alternative hypotheses. Here we assume that $\vartheta(x, y)$ can be formulated as follows:

(i) *The odds ratio function can be parameterized as $\vartheta(x, y) = \theta_\alpha\{\eta(x, y)\}$, where α is a parameter and $\eta(x, y)$ is an unspecified nuisance function.*

(ii) *For each fixed η, $\theta_\alpha(\eta)$ is a continuously differentiable function of α and $\lim_{\alpha\to\alpha_0} \theta_\alpha(\eta) = 1$, where α_0 is the parameter value under quasi-independence.*

Suppose that $\eta(x, y)$ can be estimated by $\hat{\eta}(x, y)$. Under a working assumption of independence among different tables of (x, y) and ignoring the distributions of the marginal counts, we can construct the following conditional likelihood function:

$$L(\alpha) = \prod_{x\le y} \left[\frac{\theta_\alpha\{\hat{\eta}(x, y)\}}{R(x, y) - 1 + \theta_\alpha\{\hat{\eta}(x, y)\}}\right]^{N_{11}(dx,dy)} \left[\frac{R(x, y) - 1}{R(x, y) - 1 + \theta_\alpha\{\hat{\eta}(x, y)\}}\right]^{1-N_{11}(dx,dy)}. \quad (8)$$

The corresponding score function becomes

$$\frac{\partial \log L(\alpha)}{\partial\alpha} = \iint_{x\le y} \frac{\dot{\theta}_\alpha\{\hat{\eta}(x, y)\}}{\theta_\alpha\{\hat{\eta}(x, y)\}} \left\{N_{11}(dx, dy) - \frac{N_{1\bullet}(dx, y)N_{\bullet 1}(x, dy)\theta_\alpha\{\hat{\eta}(x, y)\}}{R(x, y) - 1 + \theta_\alpha\{\hat{\eta}(x, y)\}}\right\}, \quad (9)$$

where $\dot{\theta}_\alpha(v) = \partial\theta_\alpha(v)/\partial\alpha$. Note that Eq. (8) was motivated by Clayton [15] and Oakes [16] who considered the Clayton model for bivariate censored data.

By setting $\alpha \to \alpha_0$, the score test statistic can be obtained based on Eq. (9). Specifically since $\lim_{\alpha\to\alpha_0} \theta_\alpha\{\eta(x, y)\} = 1$, the proposed score statistics has the form of L_W with the weight function

$$W(x, y) = \lim_{\alpha\to\alpha_0} \dot{\theta}_\alpha\{\hat{\eta}(x, y)\}. \quad (10)$$

Eq. (10) provides a clear guideline for choosing the weight function for L_W when the assumptions on $\vartheta(x, y)$ stated in (i) and (ii) are satisfied. The level of power improvement depends on whether $\theta_\alpha(\cdot)$ is correctly specified and how accurate $\eta(x, y)$ can be estimated. We will discuss these issues via specific examples in Section 3.2.

3.2. Semi-survival Archimedean copula models

For dependent truncation data, Chaieb, et al. [10] proposed "semi-survival" Archimedean copula (AC) models of the form

$$\pi(x, y) = \Pr(X \le x, Y > y | X \le Y) = \phi_\alpha^{-1}[\phi_\alpha\{F_X(x)\} + \phi_\alpha\{S_Y(y)\}]/c, \quad (11)$$

where c is a normalizing constant satisfying $1 = -\iint_{x\le y} d\pi(x, y)$. AC models are characterized by the generating function $\phi_\alpha(\cdot) : [0, 1] \to [0, \infty]$, where $\phi_\alpha(1) = 0$, $\phi'_\alpha(t) = \partial\phi_\alpha(t)/\partial t < 0$ and $\phi''_\alpha(t) = \partial^2\phi_\alpha(t)/\partial t^2 > 0$. Furthermore, they showed that under (11), the odds ratio function can be written as $\vartheta(x, y) = \theta_\alpha\{c\pi(x, y)\}$, where

$$\theta_\alpha(\eta) = -\eta\frac{\phi''_\alpha(\eta)}{\phi'_\alpha(\eta)}. \quad (12)$$

Hence AC models satisfy assumption (i) such that $\eta(x, y) = c\pi(x, y)$. The case of quasi-independence corresponds to $\phi(t) = -\log(t)$ in (11). After appropriate parameterization for α, we may assume that $\phi_{\alpha_0}(t) = -\log(t)$ for $\alpha_0 = 1$ so that assumption (ii) holds.

An estimator of c may be obtained using the proposal by Chaieb, et al. [10]. Alternatively, considering that $\eta(x, y) = c\pi(x, y)$ in (10) is evaluated at $\alpha \to \alpha_0$, it suffices to estimate $c = c_0$, the value under H_0. He and Yang [19] proposed to estimate $\Pr(X \le Y)$ under independence between X and Y. Although in the present case, c_0 is not necessary equivalent to $\Pr(X \le Y)$, their idea can be modified. Specifically, one can set $c_0 = F_X(X_{(1)})/\pi(X_{(1)}, X_{(1)})$ in (1) under the assumption of $S_Y(X_{(1)}) = 1$ where $X_{(1)} = \min_j X_j$. By applying the nonparametric estimator $\hat{F}_X$ of Wang, Jewell, and Tsai [7], we have $\hat{c}_0 = \hat{F}_X(X_{(1)})/\hat{\pi}(X_{(1)}, X_{(1)}-)$. Note that the same estimator $\hat{c}_0$ can also be obtained as a solution of Eq. (12) of Chaieb, Rivest and Abdous [10] by setting $\alpha = \alpha_0$ and $t = x_{(1)}$.

Now we derive the suggested form of weight in (10) for selected AC models.

Example 1 (*Clayton Copula*). The Clayton model [15] has the generating function $\phi_\alpha(t) = (t^{-(\alpha-1)} - 1)/(\alpha - 1)$ for $0 < \alpha < \infty, \alpha \ne 1$, and $\phi_{\alpha_0}(t) = -\log(t)$ when $\alpha_0 = 1$. It follows that $\theta_\alpha(\eta) = \alpha$ and hence

$$\lim_{\alpha\to\alpha_0} \dot{\theta}_\alpha\{\eta(x, y)\} = 1,$$

which corresponds to $L_{\rho=0}$, a special case of L_ρ in (4). Notice that no nuisance parameter is involved in the weight function.

Example 2 (*Frank Copula*). For Frank's model [20], the generating function has the form $\phi_\alpha(t) = \log\{(1-\alpha)/(1-\alpha^t)\}$ for $0 < \alpha < \infty, \alpha \neq 1$ and $\phi_{\alpha_0}(t) = -\log(t)$ for $\alpha_0 = 1$. Since $\theta_\alpha(\eta) = \eta\log(\alpha)/\{e^{\eta\log(\alpha)} - 1\}$ and $\theta_{\alpha_0}(\eta) = 1$, we have

$$\lim_{\alpha\to\alpha_0} \dot\theta_\alpha(\eta) = \lim_{h\to 0}\frac{1}{h}\left\{\frac{\eta\log(1+h)}{e^{\eta\log(1+h)}-1}\right\} = -\frac{\eta}{2}.$$

Thus, the suggested weight function is given by

$$\lim_{\alpha\to\alpha_0} \dot\theta_\alpha\{\eta(x,y)\} = -\frac{c\pi(x,y)}{2} \propto \pi(x,y).$$

If we estimate $\pi(x,y)$ by $\hat\pi(x,y-)$, the resulting score test becomes $L_{\rho=1}$ in (7) which is equivalent to K_a considered by Tsai [9] and Martin and Betensky [11]. This implies that these two tests are suitable for Frank's alternative.

Example 3 (*Gumbel Copula*). For the Gumbel model, the generating function equals $\phi_\alpha(t) = \{-\log(t)\}^\alpha$ for $\alpha > 1$ and $\phi_{\alpha_0}(t) = -\log(t)$ for $\alpha_0 = 1$. Under the Gumbel model, (X, Y) only permit negative association. Since $\theta_\alpha(\eta) = 1-(\alpha-1)/\log(\eta)$, it follows that

$$\lim_{\alpha\to\alpha_0} \dot\theta_\alpha\{\eta(x,y)\} \propto -1/\log\{c\pi(x,y)\}.$$

By plugging in the estimators of $\pi(x,y)$ and c in the suggested weight, we denote the corresponding test as $L_{inv\,\log}$, which however is not a member of L_ρ in (4).

In practice there may be several model choices under consideration. We suggest a heuristic approach by choosing the model that yields the highest value of $L(\hat\alpha)$, where $\hat\alpha$ maximizes $L(\alpha)$ over the corresponding parameter space. The influence of weight on the power of the corresponding test will be evaluated later via simulations.

4. Asymptotic analysis

4.1. Asymptotic normality

In this section, we state the main theoretical results. We assume that the underlying distribution is absolutely continuous under the null hypothesis in (1). Consider a class of weighted log-rank type statistics of the form,

$$L_w = \iint_{x\le y} w\{\hat\pi(x,y-)\}\left\{N_{11}(dx,dy) - \frac{N_{1\bullet}(dx,y)N_{\bullet 1}(x,dy)}{R(x,y)}\right\},$$

where $w(v)$ is a known continuously differentiable function of $v \in (0,1)$.

Theorem 1. *Under H_0, the statistics $n^{-1/2}L_w$ converges in distribution to a mean-zero normal random variable. The special case $n^{-1/2}L_\rho$ has asymptotic variance $\sigma_\rho^2 = E\{U_\rho(X_j,Y_j)^2\}$, where*

$$\begin{aligned}U_\rho(X_j,Y_j) = {} & (\rho-1)/2 \iint\iint_{x\vee x^*\le y\wedge y^*} \pi(x\vee x^*, y\wedge y^*-)^{\rho-2}\{I(X_j\le x\vee x^*, Y_j\ge y\wedge y^*) \\ & - \pi(x\vee x^*, y\wedge y^*-)\}\mathrm{sgn}\{(x-x^*)(y-y^*)\}d\pi(x,y)d\pi(x^*,y^*) \\ & - \iint\iint_{x\vee x^*\le y\wedge y^*} \pi(x\vee x^*, y\wedge y^*-)^{\rho-1}\mathrm{sgn}\{(x-x^*)(y-y^*)\} \\ & \times\{I(X_i = x, Y_i = y) + d\pi(x,y)\}d\pi(x^*,y^*). \end{aligned} \quad (13)$$

Sketch of the proofs are given in Appendices A.1 and A.2 and more complete discussions can be found in Emura and Wang [21].

4.2. Variance estimation

Eq. (7) shows that, in the absence of ties, $L_{\rho=1}$ is equivalent to K_a. Variance estimation of K_a has been discussed in Tsai [9] and Martin and Betensky [11]. Here we propose a different approach. Based on the formula in (13), we can estimate σ_ρ^2 by applying the method of moment and the plug-in principle. The arguments in Appendix A.2 yield the following variance formula for L_ρ:

$$\begin{aligned} n\hat\sigma_\rho^2 = \sum_j\Bigg[& \frac{1}{n}\sum_k I\{A_{jk}\}\hat\pi(\breve X_{jk}, \tilde Y_{jk}-)^{\rho-1}\mathrm{sgn}\{(X_j-X_k)(Y_j-Y_k)\} + \frac{(\rho+1)L_\rho}{n} \\ & + \frac{\rho-1}{n^2}\sum_{k<l} I\{A_{kl}\}\hat\pi(\breve X_{kl}, \tilde Y_{kl}-)^{\rho-2}\mathrm{sgn}\{(X_k-X_l)(Y_k-Y_l)\}I(X_j\le \breve X_{kl}, Y_j\ge \tilde Y_{kl})\Bigg]^2. \end{aligned} \quad (14)$$

This estimator incorporates the variability of estimating the nuisance function $\pi(x,y)$.

When censoring is present, analytic expressions of σ_ρ^2 become complicated and not tractable. Under general situations, the jackknife method provides a convenient tool for variance estimation. For an arbitrary weight function, the variance of L_W can be estimated by the following jackknife estimator:

$$\hat{\sigma}_{Jack}^2 = n/(n-1)\sum_j (L_W^{(-j)} - L_W^{(\cdot)})^2,$$

where $L_W^{(-j)}$ is the statistics L_W ignoring the jth observation and $L_W^{(\cdot)} = (1/n)\sum_j L_W^{(-j)}$. Emura and Wang [21] provide simulation results which compare the two variance estimators under L_ρ statistics. It is found that although the analytic estimator sometimes has better performance in variance estimation by producing smaller mean-squared errors, it tends to yield less accurate type-I probability compared with the jackknife estimator. It seems that the higher-order terms omitted in the linear expression of L_ρ still play some role in estimating the variance for finite samples.

The validity of the jackknife estimator is closely related to the smoothness of the L_ρ with respect to the empirical process $\hat{\pi}(x, y) = \sum_j I(X_j \le x, Y_j > y)/n$. This property requires a stringent smoothness condition on the corresponding statistical functional. The following theorem can be proved by checking a sufficient condition of continuous Gateaux differentiability [22] for consistency of the jackknife method.

Theorem 2. *Under H_0, the asymptotic variances σ_ρ^2 of L_ρ can be consistently estimated by the jackknife method.*

The detailed proof is given in Emura and Wang [21].

4.3. Asymptotic efficiency of the score test

The conditional likelihood constructed in Section 3 is not the true likelihood since it ignores the dependence among different tables and involves extra estimation of the nuisance parameter. Here we investigate its asymptotic efficiency. Under assumptions (i) and (ii) of Section 3.1, a Taylor series expansion of $\theta_\alpha(\eta)$ around α_0 leads to the contiguous alternative

$$H_n : \theta_{\alpha_0+n^{-1/2}}\{\eta(x, y)\} = 1 + n^{-1/2}\dot{\theta}_{\alpha_0}\{\eta(x, y)\} + o(n^{-1/2}).$$

Under the sequence of alternatives, it can be shown that the statistics $n^{-1/2}L_\rho$ converges in distribution to the normal distribution with mean

$$\mu_\rho = \lim_{n\to\infty} \frac{1}{n^2} \iint_{x\le y} \hat{\pi}^{\rho-1}(x, y)\dot{\theta}_{\alpha_0}\{\hat{\eta}(x, y)\}N_{1\bullet}(dx, y)N_{\bullet 1}(x, dy)$$

and variance σ_ρ^2. The asymptotic efficiency of L_ρ can be studied by comparing the noncentrality parameter of the chi-square test defined as

$$\tilde{\mu}_\rho^2 = \mu_\rho^2/\sigma_\rho^2.$$

Standard Cauchy–Schwarz type argument cannot be applied to obtain the optimal choice of ρ due to the complicated variance function that involves the nuisance parameter estimates. Note that $\tilde{\mu}_\rho^2$ not only depends on the alternative structure but it also functionally depends on the marginal distributions. To investigate the efficiency of L_ρ, we compute $\tilde{\mu}_\rho^2$ when the joint distribution of (X, Y) follows the Clayton and Frank AC families with selected marginal distributions, namely exponential, uniform and chi-squared distribution. The results are depicted in Fig. 1. For a range of $\rho \in [0, 2]$, the noncentrality parameter $\tilde{\mu}_\rho^2$ is maximized at $\rho = 0$ under the Clayton model and $\rho = 1$ under the Frank model for all the chosen marginal distributions. These results indicate that among all members of the L_ρ test, the score tests $L_{\rho=0}$ and $L_{\rho=1}$ are locally most powerful under the Clayton and Frank alternatives respectively.

5. Modification for right-censoring

In this section, we modify the proposed tests to adjust for right-censoring which arises when the process of observation has to be terminated before the event of interest occurs. Consider a situation that the lifetime variable Y is right-censored by C. In the presence of truncation, how to formulate the censoring mechanism deserves some discussions. We present two different ways to include the censoring mechanism. Both settings have been considered in the literature.

Case (A) The censoring variable C is also subject to the truncation criteria. Individuals satisfying $X \le C \wedge Y$ are included in the sample and otherwise truncated.

Case (B) Censoring only affects the individuals who satisfy $X \le Y$. Accordingly it is assumed that $\Pr(X \le C) = 1$.

Independent censorship means that the censoring event is not related to the disease process. In the presence of truncation, how to formulate the assumption of independent censoring depends on which censoring mechanism is adopted. Now we discuss the assumption for each setting. Chaieb, Rivest and Abdous [10] considered the situation in Case (A) and then made the following assumption:

Assumption A. C is independent of (X, Y).

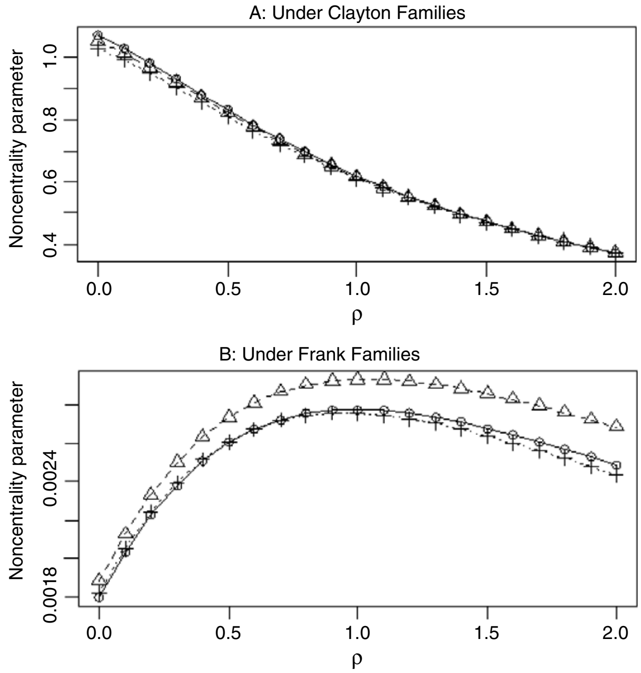

Fig. 1. Efficiency comparison of L_ρ test with $\rho \in [0, 2]$ under selected marginal distributions. ○: mean-zero exponential; △: uniform on [0,1]; +: chi-squared with one degree of freedom.

Table 2
2 × 2 table for left-truncated and right-censored data.

	$Z = y, \delta = 1$	$Z > y$	
$X = x$	$N_{11}(\mathrm{d}x, \mathrm{d}y)$		$N_{1\bullet}(\mathrm{d}x, y)$
$X < x$			
	$N_{\bullet 1}(x, \mathrm{d}y)$		$R(x, y)$

In Case (B), however, C and X cannot be independent due to the mathematical restriction $X < C$. For this case, define $C = C_R + X$, where $C_R > 0$ refers to the residual censoring time. A more proper assumption is given by

Assumption B. C_R is independent of (X, Y) given $X \le Y$.

Note that in the absence of truncation ($X = 0$ with probability one), both cases reduce to the usual independent censorship model. In the following subsections, we discuss modification of the proposed tests under the two censoring mechanisms.

5.1. The proposed test statistic under censoring

Under both censoring frameworks, observed data can be expressed as $\{(X_i, Z_i, \delta_i) : (i = 1, \ldots, n)\}$, where C_i is a random replication of C, $Z_i = Y_i \wedge C_i$ and $\delta_i = I(Y_i \le C_i)$, subject to $X_i \le Z_i$. Table 2 is a modification of Table 1 such that (x, y) denotes an uncensored failure point satisfying $x \le y$. To simplify the presentation, we use the same notations for the counts as before but modify their definitions as follows.

$$N_{11}(\mathrm{d}x, \mathrm{d}y) = \sum_j I(X_j = x, Z_j = y, \delta_j = 1), \qquad N_{1\bullet}(\mathrm{d}x, y) = \sum_j I(X_j = x, Z_j \ge y),$$

$$N_{\bullet 1}(x, \mathrm{d}y) = \sum_j I(X_j \le x, Z_j = y, \delta_j = 1) \quad \text{and} \quad R(x, y) = \sum_j I(X_j \le x, Z_j \ge y).$$

In Appendix B we show that, under H_0, the population odds ratio of Table 2 is still one under both censoring settings. Accordingly the modified log-rank statistics has the same form given below

$$L_W = \iint_{x \le y} W(x, y) \left\{ N_{11}(\mathrm{d}x, \mathrm{d}y) - \frac{N_{1\bullet}(\mathrm{d}x, y) N_{\bullet 1}(x, \mathrm{d}y)}{R(x, y)} \right\}. \tag{15}$$

Define the L_ρ statistic as

$$L_\rho = \iint_{x \le y} \hat{\upsilon}(x, y-)^\rho \left\{ N_{11}(\mathrm{d}x, \mathrm{d}y) - \frac{N_{1\bullet}(\mathrm{d}x, y) N_{\bullet 1}(x, \mathrm{d}y)}{R(x, y)} \right\}, \tag{16}$$

where $\rho \ge 0$ is a constant and $\hat{\upsilon}(x, y)$ is an estimator of $\pi(x, y)$. Note that the two censoring cases yield different consistent estimators of $\pi(x, y)$ such that

$$\hat{\upsilon}(x, y-) = \begin{cases} R(x, y)/\{n\hat{S}_C(y-)\} & \text{under Assumption A} \\ \int_0^x R(\mathrm{d}u, y)/\{n\hat{S}_{C_R}((y-u)-)\} & \text{under Assumption B}, \end{cases}$$

where $\hat{S}_C(y)$ is the product-limit estimator for $\Pr(C > y) = S_C(y)$ based on data $\{(X_i, Z_i, 1-\delta_i) \quad (i = 1, \ldots, n)\}$ [23] and $\hat{S}_{C_R}(y)$ is the usual Kaplan–Meier estimator for $\Pr(C_R > y) = S_{C_R}(y)$ based on data $\{(C_i - X_i, 1-\delta_i) : (i = 1, \ldots, n)\}$ [1]. In the absence of censoring, $\hat{\upsilon}(x, y-)$ reduces to $\hat{\pi}(x, y-)$ for both cases. Notations L_ρ (A) and L_ρ (B) will be used when $\hat{\upsilon}(x, y-)$ is defined under Assumptions A and B respectively.

Emura and Wang (2008) showed that, under Assumption A, L_ρ can be written as a Hadamard differentiable function of $\hat{H}(x, y, c) = \sum_j I(X_j \le x, Y_j > y, C_j > c)/n$:

$$L_\rho = \Phi_\rho(\hat{H}) = -\frac{n}{2} \iint \iint \iint_{x \vee x^* \le y \wedge y^* < c \wedge c^*} \frac{\{\varphi(\hat{H}; x \vee x^*, y \wedge y^* \wedge c \wedge c^*)\}^\rho}{\hat{H}\{x \vee x^*, (y \wedge y^* \wedge c \wedge c^*)-, (y \wedge y^* \wedge c \wedge c^*)-\}}$$
$$\times \operatorname{sgn}\{(x - x^*)(y - y^*)\} \mathrm{d}\hat{H}(x, y, c) \mathrm{d}\hat{H}(x^*, y^*, c^*),$$

where $\hat{\upsilon}(x, y-) \equiv \varphi(\hat{H}; x, y)$ is also a Hadamard differentiable function of $\hat{H}$. Asymptotic normality of L_ρ can be established by applying the functional delta method and the fact that $n^{1/2}(\hat{H} - H)$ converges weakly to a mean-zero Gaussian process. Emura and Wang (2008) also showed that the asymptotic variance of L_ρ can be consistently estimated by the Jackknife estimator $\hat{\sigma}^2_{Jack}$. Extension of these results under Assumption B follows essentially the same arguments by modifying the definition of $\varphi(H; x, y)$. Therefore, the test of quasi-independence can be based on $L_\rho/\hat{\sigma}_{Jack}$ by applying the asymptotic normality result.

The definition of τ_a has also been modified to account for censoring. Using the fact that the order of a pair is known for certain if the smaller one is observed, Martin and Betensky [11] define the event

$$B_{ij} = \{\breve{X}_{ij} \le \tilde{Z}_{ij}\} \cap \{(\delta_i = \delta_j = 1) \cup (Z_j - Z_i > 0 \,\&\, \delta_i = 1 \,\&\, \delta_j = 0) \cup (Z_i - Z_j > 0 \,\&\, \delta_i = 0 \,\&\, \delta_j = 1)\}$$

which is a condition for the (i, j) pairs being comparable and orderable. The modified conditional Kendall's tau, denoted as τ_b, has the same form as τ_a with A_{ij} being replaced by B_{ij}. Under quasi-independence, it can be shown that under both settings,

$$\tau_b = E[\operatorname{sgn}\{(X_i - X_j)(Z_i - Z_j)\} | B_{ij}] = 0.$$

The proof is essentially quite similar as in Appendix B and hence is omitted. In Appendix C, we show that

$$L_W = -\sum_{i<j} I\{B_{ij}\} \frac{W(\breve{X}_{ij}, \tilde{Z}_{ij})}{R(\breve{X}_{ij}, \tilde{Z}_{ij})} \operatorname{sgn}\{(X_i - X_j)(Z_i - Z_j)\}. \tag{17}$$

By setting $W(x, y) = R(x, y)/n$, L_W reduces to the empirical estimator of τ_b,

$$-\frac{1}{n} \sum_{i<j} I\{B_{ij}\} \operatorname{sgn}\{(X_i - X_j)(Z_i - Z_j)\} = -\frac{K_b}{n}.$$

Note that K_b no longer belongs to the class L_ρ in (5) when data are censored. For variance estimation, explicit variance formula for K_b was proposed by Tsai [9] based on properties of rank-statistics. Martin and Betensky [11] still apply properties of U-statistics to obtain the asymptotic variance of K_b.

5.2. Conditional score test under censoring

Now we extend the analysis in Section 3 to the two censoring settings. Extension under Assumption A is first discussed since it is more straightforward. Under the alternative hypothesis, the population odds ratio of Table 2 is $\vartheta(x, y) = \theta_\alpha\{\eta(x, y)\}$ and the arguments in Section 3.1 can be still applied based on the modified counts defined in Section 5.1. The conditional score tests is a special case of (15) with the weight function $W(x, y) = \lim_{\alpha \to \alpha_0} \dot{\theta}_\alpha\{\hat{\eta}(x, y)\}$. Consider the semi-survival AC models (11) in which $\theta_\alpha\{\eta(x, y)\}$ can be written as $\theta_\alpha(\eta) = -\eta\phi''_\alpha(\eta)/\phi'_\alpha(\eta)$ and $\eta(x, y) = c\pi(x, y)$. To estimate the nuisance parameter we rewrite it as $c\pi(x, y) = c^*\upsilon(x, y)$, where $\upsilon(x, y) = \Pr(X \le x, Z > y | X \le Z)/S_C(y)$. The nuisance parameter is estimated by $\hat{\eta}(x, y) = \hat{c}^*\hat{\upsilon}(x, y-)$, where $\hat{c}^*$ is an estimator of c^*. Under H_0, the constant c^* is

Table 3
Tests of quasi-independence for the AIDS data.

	$L_{\rho=0}$	$L_{\rho=1}$	$L_{inv\,\log}$	Tsai	M & B
Adult					
Z-value	−5.012	−2.918	−3.795	2.567	2.833
P-value	5.398×10^{-7}	3.519×10^{-3}	1.475×10^{-4}	1.027×10^{-2}	4.610×10^{-3}
$\log L(\hat{\alpha})$	−1077.878	−1080.054	−1082.860	Undefined	Undefined
Children					
Z-value	−1.838	−1.379	−1.373	0.966	1.672
P-value	0.066	0.168	0.170	0.334	0.095
$\log L(\hat{\alpha})$	−95.225	−95.434	−95.859	Undefined	Undefined

estimated by $\hat{c}_0^* = \hat{F}_X(X_{(1)})/\hat{\pi}(X_{(1)}, X_{(1)}-)$, where $\hat{F}_X$ is the estimator of Wang, Jewell, and Tsai [7] based on truncated data $\{(X_i, Z_i) : (i = 1, \ldots, n)\}$ subject to $X_i \le Z_i$. The suggested weight function under each AC model is the same as those presented in Section 3.2 except that the method of estimating nuisance parameter has to be modified. It turns out that $L_{\rho=0}$ and $L_{\rho=1}$(A) in (16) are the conditional score tests when (X, Y) follows the Clayton and Frank AC models respectively.

Derivation of the score test under Assumption B becomes more complicated since the population odds ratio of Table 2 is no longer $\theta_\alpha\{\eta(x, y)\}$. Based on (B1) of Appendix B, the odds ratio equals

$$\frac{\partial^2 \pi(x,y)/\partial x \partial y}{\partial \pi(x,y)/\partial x} \cdot \frac{\int_0^x \{\partial \pi(u,y)/\partial u\} S_{C_R}(y-u)\mathrm{d}u}{\int_0^x \{\partial^2 \pi(u,y)/\partial u \partial y\} S_{C_R}(y-u)\mathrm{d}u}.$$

This is not equal to $\theta_\alpha\{\eta(x, y)\}$ unless $S_{C_R}(y-u) = 1$ for $0 \le u \le x$. Development of the conditional score test under Assumption B will be left as our future work. Nevertheless, the choice with $W(x, y) = \lim_{\alpha \to \alpha_0} \dot{\theta}_\alpha\{\hat{\eta}(x, y)\}$ in (15) is still a valid test even it may not achieve the same level of power improvement. If censoring is light so that $\theta_\alpha\{\eta(x, y)\}$ is a good approximation of the true ratio, the resulting test will still be a good choice.

6. Numerical analysis

6.1. Data analysis

We apply the proposed methods to the aforementioned AIDS data and Channing House data and compare our results with existing analyses. The first data contains no censored observations.

Lagakos, Barraj, and De Gruttola [2] divided the AIDS data into two age groups of children (37 subjects) and adults (258 subjects) and assumed independence between the incubation time X and the lapse time Y. The Z-values and p-values of five tests are reported in Table 3. Specifically the proposed log-rank statistics based on $L_{\rho=0}$, $L_{\rho=1}$ and $L_{inv\,\log}$ utilize the jackknife method for variance estimation. The tests proposed by Tsai [9] and Martin and Betensky [11] have the form of $L_{\rho=1}$ or K_a but use their own variance estimators in the standardization. For the adult group, all the results show significant deviation from quasi-independence. The sign of the Z-values indicates positive association between X and Y ($\tau_a = 0.111$). This implies that people infected in earlier chronicle time tended to have longer length of incubation. Although similar pattern of association was also discovered in the children group ($\tau_a = 0.117$), it did not reach 5% level of statistical significance probably because the sample size is not large enough. Nevertheless H_0 is still rejected by the tests of $L_{\rho=0}$ and Martin and Betensky [11] at 10% significance level.

To determine the best weight for L_w, we compare values of the fitted likelihood under the three models, namely the Clayton, Frank and Gumbel families. In Table 3, $\log L(\hat{\alpha})$ denotes the log of conditional likelihood when $\hat{\alpha}$ is the maximized value of α over the parameter space of the model. For both covariate groups, the Clayton model is the best fitted one among the competitors and hence $L_{\rho=0}$ is recommended.

For the Channing House example, quasi-independence between a resident's lifetime (Y) and his/her entry age to the community (X) is examined under the two censoring mechanisms which differ in whether censoring could occur to a truncated subject. Six tests are compared in Table 4 which include the tests proposed by Tsai [9] and Martin and Betensky [11] and four proposed tests. The score tests, $L_{\rho=0}$, $L_{\rho=1}$(A) and $L_{inv\,\log}$, use the suggested weights for the three AC models respectively with $\hat{v}(x, y-)$ defined under Assumption A. The $L_{\rho=1}$ (B) test adopts Assumption B to define $\hat{v}(x, y-)$. All the tests are valid.

The first analysis uses the data provided in Hyde [4] which contains 462 (97 men and 365 women) subjects. Among them, 286 people withdrew from the community yielding the censoring proportion 0.62. Based on the first half of Table 4, the Z-value of each test indicates slightly positive association between X and Y ($\tau_b = 0.088$). The four tests, namely $L_{\rho=1}$ (A), $L_{\rho=1}$(B), Tsai's test and Martin and Betensky's test, reach the 10% significance level. In fact, the likelihood analysis favors the Frank model under which the score test is $L_{\rho=1}$ (A). Recall that in the presence of censoring, Tsai and

Table 4
Tests of quasi-independence for the Channing House data.

	$L_{\rho=0}$	$L_{\rho=1}$ (A)	$L_{inv\,\log}$	$L_{\rho=1}$ (B)	Tsai	M & B
(1) 462 subjects						
Z-value	−0.515	−1.669	−1.169	−1.700	1.776	1.837
P-value	0.607	0.095	0.243	0.089	0.076	0.066
$\log L(\hat{\alpha})$	−809.207	−807.954	−809.316	Undefined	Undefined	Undefined
(2) 97 men, a subset of (1)						
Z-value	−1.286	−1.379	−1.116	−1.973	2.021	2.053
P-value	0.198	0.168	0.264	0.048	0.043	0.040
$\log L(\hat{\alpha})$	−139.297	−139.267	−140.268	Undefined	Undefined	Undefined

Note: $L_{\rho=1}$ (A) uses the weight function $\hat{v}(x, y-) = R(x, y)/\{n\hat{S}_C(y-)\}$ and $L_{\rho=1}$ (B) uses the weight function $\hat{v}(x, y-) = \int_0^x R(du, y)/\{n\hat{S}_{C_R}((y-u)-)\}$.

Martin and Betensky's tests use the weight $R(x, y)/n$ while $L_{\rho=1}$ (A) and $L_{\rho=1}$ (B) adopt the weight $R(x, y)/\{n\hat{S}_C(y-)\}$ and $\int_0^x R(du, y)/\{n\hat{S}_{C_R}((y-u)-)\}$ respectively. Hence they are no longer equivalent.

The second analysis uses the data in Hyde [3], where only the 97 men were studied with 51 subjects being censored. This subset also reveals positive association between X and Y ($\tau_b = 0.199$). Based on the second half of Table 4, the three score tests fail to reject quasi-independence. The values of maximized log-likelihood still favor the Frank alternative in which the score test is $L_{\rho=1}$ (A) with the p-value 0.168. In contrast, three tests $L_{\rho=1}$ (B), Tsai's test and Martin and Betensky's test suggest rejecting quasi-independence at 5% level (p-values: 0.048, 0.043 and 0.040 respectively).

Now we discuss the results of Channing House data in more detail. Firstly, the methods of variance estimation seem to have not much effect. In fact, if we tested the second dataset using L_W with $W(x, y) = R(x, y)/n$ and the jackknife variance estimator, the corresponding Z-value becomes −2.033 (p-value: 0.042) which is very close to the results based on the two competing tests. Therefore the test result seems to be mostly affected by the chosen weight function. Note that the function $R(x, y)/n$ assigns higher weight to early failure time y than $R(x, y)/\{n\hat{S}_C(y-)\}$ in $L_{\rho=1}$(A). We suspect that the association at earlier time period is higher for the subset of men than it is for the whole sample of 462 subjects. Notice that for the male group, the $L_{\rho=1}$ (B) test rejects H_0 (p-value: 0.048) while the $L_{\rho=1}$ (A) test does not (p-value: 0.168). To determine which censoring assumption is more suitable, one may further examine whether the censoring event can also be truncated or not. For Channing House data among 286 censored subjects, 27 subjects withdrew from the study and the remaining 259 subjects survived until the end of study. The reason of withdraw might be due to financial insecurity. Assumption B (i.e. $\Pr(X \leq C) = 1$) is more plausible if the end-of-study effect was the primary reason of censoring. However if the financial issue still affected a person's decision on the enrollment of the community, Assumption A may be a proper choice. In addition the definition of the target population is crucial. The researcher might adopt Assumption B by excluding those who were not rich enough to enroll from the target population.

6.2. Simulation studies

Finite-sample performances of the proposed test and their competitors are evaluated via simulations. Random pairs of (X, Y) were generated from three well-known semi-survival AC models, namely the Clayton, Frank and Gumbel families discussed in Section 3.2 The level of association for an AC model can be described in terms of (pre-truncated) Kendall's tau defined as $\tau = E[\text{sgn}(X_i - X_j)(Y_i - Y_j)]$ which is independent of the marginal distributions. Since the major goal of the simulations is to see the power improvement in the suggested weight in (10), we adopt Assumption A for the censoring mechanism, under which the score tests are derived. Accordingly, the censoring variable C was generated independently from (X, Y). The marginals of (X, Y, C) follow exponential distributions with the hazard rates yielding the targeted levels of $c = P(X \leq Y)$ (i.e. 66.7%, 50.0% and 33.3%) for the uncensored case and of $c^* = P(X \leq Z)$ (i.e. 66.7%, 50.0% and 33.3%) for the 50% censored case ($P(C < Y|X \leq Z) = 0.5$) respectively. For each setting, we provide the value of conditional Kendall's tau τ_a or τ_b.

We consider three proposed tests, namely $L_{\rho=0}$, $L_{\rho=1}$ and $L_{inv\,\log}$, using the jackknife method for variance estimation. For the Clayton, Frank and Gumbel alternatives, the score tests correspond to $L_{\rho=0}$, $L_{\rho=1}$ and $L_{inv\,\log}$ respectively. The tests proposed by Tsai [9] and Martin and Betensky [11] are also evaluated. In the absence of censoring, these two tests constructed based on K_a are equivalent to $L_{\rho=1}$ except that different variance estimators are used. Performances of the five tests at $n = 100$ and 200 are studied.

Tables 5 and 6 summarize the results based on 500 replications when (X, Y) follow the Clayton model. Under quasi-independence, the rejection probability for all tests are close to the nominal 5% level, and as expected, the power of each test increases as the level of association departs from quasi-independence. In all the cases, the proposed score test $L_{\rho=0}$ is uniformly more powerful than the other tests. The test $L_{\rho=1}$ and two related tests proposed by Tsai [9] and Martin and Betensky [11] have similar and sometimes unsatisfactory performances. Also, the power of each test improves when the censoring rate decreases.

Table 5
Empirical rejection probabilities of the proposed tests ($L_{\rho=0}$, $L_{\rho=1}$ and $L_{inv\,\log}$) and two competing tests (Tsai's and Martin and Betensky's tests) at level $\alpha = 0.05$ based on 500 replications when (X, Y) under the Clayton model with sample size 100.

$c = \Pr(X \le Y)$ $c^* = \Pr(X \le Z)$	$\tau\,(\tau_a/\tau_b)$	$L_{\rho=0}$	$L_{\rho=1}$	$L_{inv\,\log}$	Tsai	M & B
Uncensored						
$c = 0.667$	−0.2 (−0.200)	0.908	0.832	0.860	0.856	0.800
	−0.1 (−0.100)	0.410	0.320	0.334	0.344	0.312
	0.0 (0.000)	0.052	0.044	0.042	0.046	0.046
	0.1 (0.100)	0.518	0.374	0.442	0.358	0.378
	0.2 (0.200)	0.998	0.914	0.962	0.900	0.910
$c = 0.500$	−0.2 (−0.200)	0.900	0.802	0.852	0.832	0.786
	−0.1 (−0.100)	0.404	0.290	0.344	0.334	0.280
	0.0 (0.000)	0.062	0.054	0.044	0.052	0.064
	0.1 (0.100)	0.456	0.354	0.376	0.322	0.372
	0.2 (0.200)	0.998	0.912	0.984	0.888	0.914
$c = 0.333$	−0.2 (−0.200)	0.900	0.794	0.838	0.846	0.786
	−0.1 (−0.100)	0.396	0.290	0.332	0.340	0.272
	0.0 (0.000)	0.046	0.036	0.032	0.036	0.038
	0.1 (0.100)	0.518	0.382	0.438	0.352	0.410
	0.2 (0.200)	0.990	0.896	0.978	0.900	0.920
50% Censored						
$c^* = 0.667$	−0.2 (−0.200)	0.746	0.622	0.606	0.604	0.582
	−0.1 (−0.100)	0.262	0.186	0.192	0.172	0.148
	0.0 (0.000)	0.056	0.054	0.028	0.048	0.052
	0.1 (0.100)	0.222	0.212	0.198	0.184	0.182
	0.2 (0.200)	0.836	0.690	0.734	0.646	0.636
$c^* = 0.500$	−0.2 (−0.200)	0.696	0.552	0.558	0.538	0.512
	−0.1 (−0.100)	0.270	0.176	0.172	0.190	0.162
	0.0 (0.000)	0.038	0.034	0.026	0.050	0.046
	0.1 (0.100)	0.244	0.220	0.214	0.204	0.204
	0.2 (0.200)	0.824	0.660	0.702	0.622	0.624
$c^* = 0.333$	−0.2 (−0.200)	0.690	0.542	0.542	0.522	0.482
	−0.1 (−0.100)	0.254	0.158	0.156	0.166	0.140
	0.0 (0.000)	0.046	0.060	0.044	0.054	0.056
	0.1 (0.100)	0.204	0.170	0.154	0.178	0.188
	0.2 (0.200)	0.852	0.688	0.740	0.676	0.686

Note: For each run, five test statistics are calculated based on the same dataset.

The results for the Frank model under different levels of association are summarized in Table 7 ($n = 100$) and Table 8 ($n = 200$). As mentioned earlier, the score test based on $L_{\rho=1}$ and the tests proposed by Tsai [9] and Martin and Betensky [11] use the same weight function when data are not censored. Under the Frank model, the three tests have shown higher power than both $L_{\rho=0}$ and $L_{inv\,\log}$ as expected but a clear-cut dominance among the three is not found. Compared with the Clayton case, the magnitude of power improvement reduces a little bit. This may be due to the effect of estimating the nuisance function of $\pi(x, y)$ in the suggested weight for the Frank model.

Table 9 contains the results under the Gumbel model with $\tau = -0.2$ and $\tau = -0.4$ since the semi-survival Gumbel model only permits negative association. In contrast to the Clayton and Frank models, the discrepancy for the power curves of different tests becomes less clear. Nevertheless for the uncensored case with $n = 200$, the proposed score test based on $L_{inv\,\log}$ still performs slightly better than the competing tests. We suspect that the gain by using the suggested form of weight $1/\log\{c\pi(x, y)\}$ may be somewhat offset by estimating two nuisance parameters c and $\pi(x, y)$.

Interestingly the level of truncated proportion has a clear impact on the power performance if the data follow the Frank or Gumbel models, while it does not under the Clayton model. Now we provide some heuristic explanations. Under these two models, the odds ratio function $\vartheta(x, y)$ is a monotone function of $c = P(X \le Y)$ or $c^* = P(X \le Z)$. It turns out that the power of all tests increases as c or c^* gets larger. In contrast, $\vartheta(x, y) = \alpha$ under the Clayton model and this may explain why the power of the tests is not much affected by c or c^*.

In general, the simulation results confirm that the suggested weight in (10) can improve the power when the alternative is correctly specified. On the other hand, a wrong choice of weight may result in loss of power. The results of the simulation studies are consistent with the efficiency study in Section 4.3.

7. Concluding remarks

A related area of research is testing independence for bivariate failure times. Rank-based procedures were proposed by Cuzick [24,25] and Dabrowska [26]. Oakes [27] suggested a concordance test based on an estimate of Kendall's tau which keeps the information of ranks and has a nice expression as a U-statistic. Shih and Louis [28,29] utilized the covariance process of martingale residuals to construct test statistics. Hsu and Prentice [30] generalized the idea of Mantel–Haenszel statistics to test independence for right-censored data. Similar idea has been extended to bivariate current status data by Ding and Wang [31] based on another formulation of 2×2 tables.

Table 6
Empirical rejection probabilities of the proposed tests ($L_{\rho=0}$, $L_{\rho=1}$ and $L_{inv\,\log}$) and two competing tests (Tsai's and Martin and Betensky's tests) at level $\alpha = 0.05$ based on 500 replications when (X, Y) under the Clayton model with sample size 200.

$c = \Pr(X \le Y)$ $c^* = \Pr(X \le Z)$	$\tau\,(\tau_a/\tau_b)$	$L_{\rho=0}$	$L_{\rho=1}$	$L_{inv\,\log}$	Tsai	M & B
Uncensored						
$c = 0.667$	−0.2 (−0.200)	0.990	0.970	0.988	0.974	0.970
	−0.1 (−0.100)	0.706	0.534	0.626	0.590	0.522
	0.0 (0.000)	0.048	0.056	0.044	0.060	0.050
	0.1 (0.100)	0.872	0.646	0.798	0.622	0.658
	0.2 (0.200)	1.000	1.000	0.998	1.000	1.000
$c = 0.500$	−0.2 (−0.200)	1.000	0.984	0.992	0.984	0.978
	−0.1 (−0.100)	0.684	0.520	0.614	0.566	0.514
	0.0 (0.000)	0.044	0.056	0.044	0.060	0.054
	0.1 (0.100)	0.874	0.670	0.824	0.642	0.676
	0.2 (0.200)	1.000	0.998	1.000	0.998	0.998
$c = 0.333$	−0.2 (−0.200)	0.990	0.974	0.986	0.982	0.974
	−0.1 (−0.100)	0.684	0.510	0.628	0.570	0.504
	0.0 (0.000)	0.052	0.070	0.046	0.058	0.070
	0.1 (0.100)	0.886	0.678	0.822	0.642	0.696
	0.2 (0.200)	1.000	1.000	1.000	0.998	1.000
50% Censored						
$c^* = 0.667$	−0.2 (−0.200)	0.936	0.854	0.868	0.842	0.838
	−0.1 (−0.100)	0.504	0.378	0.384	0.340	0.316
	0.0 (0.000)	0.048	0.052	0.048	0.054	0.046
	0.1 (0.100)	0.488	0.376	0.394	0.324	0.328
	0.2 (0.200)	0.996	0.944	0.974	0.910	0.908
$c^* = 0.500$	−0.2 (−0.200)	0.940	0.866	0.894	0.840	0.828
	−0.1 (−0.100)	0.456	0.348	0.360	0.332	0.308
	0.0 (0.000)	0.062	0.050	0.058	0.054	0.060
	0.1 (0.100)	0.566	0.446	0.494	0.394	0.408
	0.2 (0.200)	0.992	0.938	0.978	0.924	0.926
$c^* = 0.333$	−0.2 (−0.200)	0.920	0.834	0.872	0.834	0.824
	−0.1 (−0.100)	0.450	0.354	0.380	0.338	0.330
	0.0 (0.000)	0.044	0.046	0.048	0.052	0.042
	0.1 (0.100)	0.500	0.382	0.446	0.340	0.364
	0.2 (0.200)	0.998	0.946	0.984	0.956	0.960

Note: For each run, five test statistics are calculated based on the same dataset.

This article considers left-truncated data in the presence of right-censoring. A modified version of Kendall's tau was proposed by Tsai [9] and then used as the basis for testing quasi-independence by both Tsai [9] and Martin and Betensky [11]. Alternatively we apply the idea of log-rank type statistics based on 2×2 tables designed for describing truncation data. By permitting a flexible weight function, the proposed statistics form a general class of tests. A nice equivalence property between the log-rank type statistics and Kendall's tau statistics has been established. This relationship allows us to compare different types of tests under a unified framework and it turns out that the weight function plays a crucial role. The distributional properties of the 2×2 tables shed some light on the underlying likelihood structure. Accordingly, motivated by the papers of Clayton [15] and Oakes [16], we derive a score test when the dependence structure under the alternative hypothesis can be modeled via the odds ratio function $\vartheta(x, y)$. Compared with the conditional Kendall's tau measures, $\vartheta(x, y)$ is a better association measure since it is independent of the marginal distributions and can be accurately estimated in the presence of censoring. The proposed score test has the log-rank type expression with the weight function chosen to fit the alternative hypothesis and hence has good power when the true model is assumed. The functional delta method is applied to derive large-sample properties for the proposed test statistics with flexible weight functions which may contain nuisance parameters. Consistency of the jackknife variance estimator is also justified.

To find the score test, a heuristic model selection procedure is proposed by comparing the values of the conditional likelihood functions under different model choices. Alternatively Beaudoin and Lakhal-Chaieb [32] proposed a different method for model selection. They also suggested fitting the AIDS data by the Clayton model and Channing House data by the Frank model.

In the analysis of the Channing House data, we discuss the issue of choosing a suitable assumption on censoring. In summary, one should check whether the reason of censoring can occur to those with $X > Y$. This assumption also depends on how the target population is defined.

For analyzing more complicated truncation and censoring structures, Martin and Betensky [11] considered several extended versions of Kendall's tau and utilized properties of U-statistics in variance estimation and large-sample analysis. It would be interesting to apply the idea of log-rank tests to these data settings. This extension is not trivial since the formulation of appropriate "risk sets" in the construction of 2×2 tables is not straightforward. We will leave this problem as a future research topic.

Table 7
Empirical rejection probabilities of the proposed tests ($L_{\rho=0}$, $L_{\rho=1}$ and $L_{inv\,\log}$) and two competing tests (Tsai's and Martin and Betensky's tests) at level $\alpha = 0.05$ based on 500 replications when (X, Y) under Frank's model with sample size 100.

$c = \Pr(X \le Y)$ $c^* = \Pr(X \le Z)$	$\tau\,(\tau_a/\tau_b)$	$L_{\rho=0}$	$L_{\rho=1}$	$L_{inv\,\log}$	Tsai	M & B
Uncensored						
$c = 0.667$	−0.4 (−0.242)	0.864	0.956	0.946	0.956	0.952
	−0.2 (−0.103)	0.292	0.348	0.322	0.366	0.330
	0.0 (0.000)	0.052	0.044	0.042	0.046	0.046
	0.2 (0.081)	0.214	0.256	0.236	0.234	0.270
	0.4 (0.163)	0.532	0.738	0.620	0.722	0.742
$c = 0.500$	−0.4 (−0.189)	0.664	0.852	0.806	0.848	0.830
	−0.2 (−0.075)	0.166	0.206	0.198	0.216	0.182
	0.0 (0.000)	0.062	0.054	0.044	0.052	0.064
	0.2 (0.047)	0.114	0.126	0.102	0.116	0.152
	0.4 (0.082)	0.216	0.234	0.208	0.244	0.286
$c = 0.333$	−0.4 (−0.135)	0.406	0.544	0.498	0.552	0.510
	−0.2 (−0.050)	0.130	0.132	0.082	0.142	0.126
	0.0 (0.000)	0.046	0.036	0.032	0.036	0.038
	0.2 (0.026)	0.060	0.084	0.046	0.076	0.090
	0.4 (0.034)	0.064	0.066	0.056	0.062	0.106
50% Censored						
$c^* = 0.667$	−0.4 (−0.340)	0.856	0.926	0.912	0.930	0.908
	−0.2 (−0.141)	0.310	0.346	0.320	0.334	0.314
	0.0 (0.000)	0.052	0.044	0.042	0.046	0.046
	0.2 (0.131)	0.256	0.338	0.278	0.324	0.316
	0.4 (0.284)	0.752	0.898	0.834	0.888	0.884
$c^* = 0.500$	−0.4 (−0.260)	0.720	0.806	0.782	0.786	0.772
	−0.2 (−0.131)	0.224	0.230	0.206	0.244	0.216
	0.0 (0.000)	0.062	0.054	0.044	0.052	0.064
	0.2 (0.080)	0.144	0.190	0.146	0.174	0.186
	0.4 (0.167)	0.456	0.568	0.468	0.550	0.564
$c^* = 0.333$	−0.4 (−0.223)	0.474	0.490	0.444	0.518	0.462
	−0.2 (−0.081)	0.142	0.126	0.108	0.146	0.124
	0.0 (0.000)	0.046	0.036	0.032	0.036	0.038
	0.2 (0.049)	0.062	0.082	0.046	0.084	0.100
	0.4 (0.097)	0.122	0.148	0.104	0.178	0.196

Note: For each run, five test statistics are calculated based on the same dataset.

Acknowledgment

The second author's research was financially supported by the research grant 95-2118-M-009-005 funded by National Science Council of Taiwan, ROC.

Appendix A. Asymptotic analysis

Let $D\{[0,\infty)^2\}$ be the collection of all right-continuous functions with left-side limit defined on $[0,\infty)^2$, whose norm is defined by $\|f(x,y)\|_\infty = \sup_{x,y}|f(x,y)|$ for $f \in D\{[0,\infty)^2\}$. We assume that the function $\pi(x,y) = F_X(x)S_Y(y)/c_0$ is absolutely continuous. The empirical process on the plane is defined as:

$$\hat{\pi}(x,y) = \frac{1}{n}\sum_{j=1}^{n} I(X_j \le x, Y_j > y).$$

The functional delta method is applied based on the weak convergence result of $n^{1/2}(\hat{\pi}(x,y) - \pi(x,y))$ to a mean 0 Gaussian process $V(x,y)$ on $D\{[0,\infty)^2\}$ with the covariance structure given by

$$cov\{V(x_1,y_1), V(x_2,y_2)\} = \pi(x_1 \wedge x_2, y_1 \vee y_2) - \pi(x_1,y_1)\pi(x_2,y_2),$$

for any $(x_1,y_1), (x_2,y_2) \in [0,\infty)^2$.

A.1. *Proof of* Theorem 1

After some algebraic manipulations based on (6), we can write

$$L_w = -\frac{1}{2n}\sum_{i,j} I\{A_{ij}\}\frac{w\{\hat{\pi}(\breve{X}_{ij}, \tilde{Y}_{ij}-)\}}{\hat{\pi}(\breve{X}_{ij}, \tilde{Y}_{ij}-)}\mathrm{sgn}\{(X_i - X_j)(Y_i - Y_j)\}.$$

Table 8
Empirical rejection probabilities of the proposed tests ($L_{\rho=0}$, $L_{\rho=1}$ and $L_{inv\,\log}$) and two competing tests (Tsai's and Martin and Betensky's tests) at level $\alpha = 0.05$ based on 500 replications when (X, Y) under Frank's model with sample size 200.

$c = \Pr(X \le Y)$ $c^* = \Pr(X \le Z)$	$\tau\,(\tau_a/\tau_b)$	$L_{\rho=0}$	$L_{\rho=1}$	$L_{inv\,\log}$	Tsai	M & B
Uncensored						
$c = 0.667$	−0.4 (−0.242)	0.990	1.000	0.996	1.000	1.000
	−0.2 (−0.103)	0.456	0.596	0.576	0.606	0.600
	0.0 (0.000)	0.048	0.056	0.044	0.060	0.050
	0.2 (0.081)	0.340	0.486	0.406	0.452	0.490
	0.4 (0.163)	0.820	0.970	0.906	0.968	0.968
$c = 0.500$	−0.4 (−0.189)	0.912	0.972	0.966	0.970	0.972
	−0.2 (−0.075)	0.260	0.410	0.370	0.428	0.418
	0.0 (0.000)	0.044	0.056	0.044	0.060	0.054
	0.2 (0.047)	0.188	0.242	0.208	0.226	0.236
	0.4 (0.082)	0.334	0.434	0.358	0.438	0.468
$c = 0.333$	−0.4 (−0.135)	0.654	0.812	0.792	0.824	0.810
	−0.2 (−0.050)	0.138	0.158	0.150	0.166	0.160
	0.0 (0.000)	0.052	0.070	0.046	0.058	0.070
	0.2 (0.026)	0.072	0.086	0.064	0.092	0.088
	0.4 (0.034)	0.104	0.148	0.102	0.148	0.180
50% Censored						
$c^* = 0.667$	−0.4 (−0.340)	0.990	0.992	0.996	1.000	1.000
	−0.2 (−0.141)	0.530	0.620	0.596	0.608	0.576
	0.0 (0.000)	0.048	0.056	0.044	0.060	0.050
	0.2 (0.131)	0.426	0.582	0.528	0.562	0.574
	0.4 (0.284)	0.984	0.984	0.990	1.000	1.000
$c^* = 0.500$	−0.4 (−0.260)	0.946	0.976	0.976	0.978	0.974
	−0.2 (−0.131)	0.338	0.392	0.382	0.410	0.398
	0.0 (0.000)	0.044	0.056	0.044	0.060	0.054
	0.2 (0.080)	0.236	0.344	0.298	0.328	0.336
	0.4 (0.167)	0.724	0.862	0.772	0.844	0.848
$c^* = 0.333$	−0.4 (−0.223)	0.734	0.806	0.792	0.804	0.780
	−0.2 (−0.081)	0.216	0.212	0.202	0.220	0.196
	0.0 (0.000)	0.052	0.070	0.046	0.058	0.070
	0.2 (0.049)	0.106	0.128	0.102	0.136	0.136
	0.4 (0.097)	0.248	0.334	0.232	0.290	0.318

Note: For each run, five test statistics are calculated based on the same dataset.

This allows us to rewrite the statistics as

$$\begin{aligned} L_w &= -\frac{n}{2}\iint\iint_{x\vee x^*\le y\wedge y^*}\frac{w\{\hat\pi(x\vee x^*, y\wedge y^*-)\}}{\hat\pi(x\vee x^*, y\wedge y^*-)}\mathrm{sgn}\{(x-x^*)(y-y^*)\}\mathrm{d}\hat\pi(x,y)\mathrm{d}\hat\pi(x^*,y^*) \\ &\equiv -n\Phi(\hat\pi), \end{aligned}$$

where the definition of the functional $\Phi(\cdot) : D\{[0,\infty)^2\} \to \mathbf{R}$ is

$$\Phi(\pi) = \iint\iint_{x\vee x^*\le y\wedge y^*}\frac{w\{\pi(x\vee x^*, y\wedge y^*-)\}}{2\pi(x\vee x^*, y\wedge y^*-)}\mathrm{sgn}\{(x-x^*)(y-y^*)\}\mathrm{d}\pi(x,y)\mathrm{d}\pi(x^*,y^*).$$

By setting the argument π as $\pi(x,y) = \Pr(X \le x, Y > y|X \le Y)$ and viewing the above integral as an expectation, we have $\Phi(\pi) = 0$:

$$\begin{aligned} \Phi(\pi) &= E\left[I\{A_{12}\}\frac{w\{\pi(\breve X_{12}, \tilde Y_{12}-)\}}{2}\pi(\breve X_{12}, \tilde Y_{12}-)\mathrm{sgn}\{(X_1-X_2)(Y_1-Y_2)\}\right] \\ &= E\left[I\{A_{12}\}\frac{w\{\pi(\breve X_{12}, \tilde Y_{12}-)\}}{2}\pi(\breve X_{12}, \tilde Y_{12}-)E\{\mathrm{sgn}\{(X_1-X_2)(Y_1-Y_2)\}|\breve X_{12}, \tilde Y_{12}\}\right] = 0. \end{aligned}$$

By direct calculations, we can show the Hadamard differentiability of $\Phi(\cdot)$. The differential map of $\Phi(\cdot)$ at $\pi \in D\{[0,\infty)^2\}$ with direction $h \in D\{[0,\infty)^2\}$ is

$$\begin{aligned} \Phi'_\pi(h) &= \iint\iint_{x\vee x^*\le y\wedge y^*}\frac{w'\{\pi(x\vee x^*, y\wedge y^*-)\}}{2\pi(x\vee x^*, y\wedge y^*-)}h(x\vee x^*, y\wedge y^*-)\mathrm{sgn}\{(x-x^*)(y-y^*)\}\mathrm{d}\pi(x,y)\mathrm{d}\pi(x^*,y^*) \\ &\quad - \iint\iint_{x\vee x^*\le y\wedge y^*}\frac{w\{\pi(x\vee x^*, y\wedge y^*-)\}}{2\pi(x\vee x^*, y\wedge y^*-)^2}h(x\vee x^*, y\wedge y^*-)\mathrm{sgn}\{(x-x^*)(y-y^*)\}\mathrm{d}\pi(x,y)\mathrm{d}\pi(x^*,y^*) \\ &\quad + \iint\iint_{x\vee x^*\le y\wedge y^*}\frac{w\{\pi(x\vee x^*, y\wedge y^*-)\}}{\pi(x\vee x^*, y\wedge y^*-)}\mathrm{sgn}\{(x-x^*)(y-y^*)\}\mathrm{d}h(x,y)\mathrm{d}\pi(x^*,y^*). \end{aligned}$$

Table 9
Empirical rejection probabilities of three proposed tests ($L_{\rho=0}$, $L_{\rho=1}$ and $L_{inv\,\log}$) and two competing tests (Tsai's and Martin and Betensky's tests) at level $\alpha = 0.05$ based on 500 replications when (X, Y) under Gumbel's model with sample sizes 100 and 200.

$c = \Pr(X \le Y)$ $c^* = \Pr(X \le Z)$	$\tau(\tau_a/\tau_b)$	$L_{\rho=0}$	$L_{\rho=1}$	$L_{inv\,\log}$	Tsai	M & B
$n = 100$, uncensored						
$c = 0.67$	−0.4 (−0.200)	0.804	0.844	0.836	0.860	0.820
	−0.2 (−0.081)	0.226	0.224	0.220	0.234	0.206
$c = 0.50$	−0.4 (−0.169)	0.712	0.722	0.730	0.756	0.700
	−0.2 (−0.063)	0.196	0.174	0.168	0.186	0.168
$c = 0.33$	−0.4 (−0.138)	0.552	0.478	0.500	0.524	0.466
	−0.2 (−0.054)	0.142	0.124	0.116	0.126	0.116
$n = 100$, 50% censored						
$c^* = 0.67$	−0.4 (−0.293)	0.766	0.804	0.784	0.808	0.774
	−0.2 (−0.119)	0.234	0.204	0.190	0.206	0.184
$c^* = 0.50$	−0.4 (−0.215)	0.658	0.670	0.670	0.668	0.626
	−0.2 (−0.085)	0.134	0.146	0.126	0.144	0.134
$c^* = 0.33$	−0.4 (−0.198)	0.504	0.418	0.418	0.420	0.408
	−0.2 (−0.059)	0.134	0.092	0.092	0.088	0.078
$n = 200$, uncensored						
$c = 0.67$	−0.4 (−0.200)	0.978	0.992	0.994	0.990	0.990
	−0.2 (−0.081)	0.360	0.376	0.392	0.386	0.364
$c = 0.50$	−0.4 (−0.169)	0.934	0.944	0.950	0.950	0.936
	−0.2 (−0.063)	0.308	0.302	0.312	0.310	0.306
$c = 0.33$	−0.4 (−0.138)	0.828	0.792	0.830	0.824	0.786
	−0.2 (−0.054)	0.226	0.200	0.208	0.212	0.204
$n = 200$, 50% censored						
$c^* = 0.67$	−0.4 (−0.293)	0.960	0.978	0.980	0.976	0.976
	−0.2 (−0.119)	0.356	0.380	0.380	0.394	0.360
$c^* = 0.50$	−0.4 (−0.215)	0.878	0.884	0.890	0.880	0.868
	−0.2 (−0.085)	0.262	0.246	0.246	0.234	0.226
$c^* = 0.33$	−0.4 (−0.198)	0.722	0.684	0.696	0.692	0.674
	−0.2 (−0.059)	0.186	0.158	0.162	0.168	0.158

Note: For each run, five test statistics are calculated based on the same dataset.

Applying the functional delta method [33], we obtain the asymptotic expression

$$n^{-1/2}L_w = -n^{1/2}\Phi(\hat{\pi}) = -n^{-1/2}\sum_j \Phi'_\pi(\delta_{(X_j,Y_j)} - \pi) + o_P(1),$$

where $\delta_{(X_j,Y_j)}(x, y) = I(X_j \le x, Y_j > y)$. It is easy to see that the sequences,

$$U(X_j, Y_j) \equiv \Phi'_\pi(\delta_{(X_j,Y_j)} - \pi) \quad \text{for } j = 1, \ldots, n,$$

are *iid* random variables with mean-zero. From the central limit theorem, $n^{-1/2}L_w$ converges to a mean-zero normal distribution with the variance $\sigma^2 = E[U(X_j, Y_j)^2]$.

A.2. Analytic variance estimator for the G^ρ class

Recall that the L_ρ class is a sub-family of L_w. For this class, one can obtain the explicit formula of $U_\rho(X_j, Y_j)$ given in (13). Accordingly it is not difficult to obtain an analytic estimator of σ^2 based on (13) as follows: The derivative map is given by

$$\Phi'_\pi(h) = (\rho - 1)/2 \iint \iint_{x\vee x*\le y\wedge y*} \pi(x \vee x^*, y \wedge y^*-)^{\rho-2} h(x \vee x^*, y \wedge y^*-) \times \operatorname{sgn}\{(x - x^*)(y - y^*)\} d\pi(x, y) d\pi(x^*, y^*) + \iint \iint_{x\vee x*\le y\wedge y*} \pi(x \vee x^*, y \wedge y^*-)^{\rho-1} \operatorname{sgn}\{(x - x^*)(y - y^*)\} dh(x, y) d\pi(x^*, y^*).$$

Hence the asymptotic variance of L_ρ can be estimated by $\sum_j \Phi'_{\hat{\pi}}(\delta_{(X_j,Y_j)} - \hat{\pi})^2$, where

$$\Phi'_{\hat{\pi}}(\delta_{(X_j,Y_j)} - \hat{\pi}) = (\rho - 1)/2 \iint \iint_{x\vee x*\le y\wedge y*} \hat{\pi}(x \vee x*, y \wedge y*-)^{\rho-2} \times \left\{ I(X_j \le x \vee x*, Y_j \ge y \wedge y*) - \hat{\pi}(x \vee x*, y \wedge y*-) \right\}$$

$$\times \operatorname{sgn}\{(x-x*)(y-y*)\}\mathrm{d}\hat{\pi}(x,y)\mathrm{d}\hat{\pi}(x*,y*)$$
$$-\iint\iint_{x\vee x*\le y\wedge y*}\hat{\pi}(x\vee x*,y\wedge y*-)^{\rho-1}\operatorname{sgn}\{(x-x*)(y-y*)\}$$
$$\times\left\{I(X_j=x,Y_j=y)+\mathrm{d}\hat{\pi}(x,y)\right\}\mathrm{d}\hat{\pi}(x*,y*)$$
$$=\frac{1}{n}\sum_k I\{A_{jk}\}\hat{\pi}(\check{X}_{jk},\tilde{Y}_{jk}-)^{\rho-1}\operatorname{sgn}\{(X_j-X_k)(Y_j-Y_k)\}+\frac{(\rho+1)L_\rho}{n}$$
$$+\frac{\rho-1}{n^2}\sum_{k<l}I\{A_{kl}\}\hat{\pi}(\check{X}_{kl},\tilde{Y}_{kl}-)^{\rho-2}\operatorname{sgn}\{(X_k-X_l)(Y_k-Y_l)\}I(X_j\le\check{X}_{kl},Y_j\ge\tilde{Y}_{kl}).$$

Based on the above expression, one can estimate the asymptotic variance $AVar(L_\rho)=n\sigma_\rho^2$ by Eq. (14).

Appendix B. Odds ratio of Table 2

Assume that all the time variables are continuous. Under H_0 and Assumption A, all entries in Table 2 is observed under the conditioning event $X\le Z$. Thus, the population odds ratio of Table 2 can be written as

$$\frac{\Pr(X=x,Z=y,\delta=1|X\le Z)}{\Pr(X=x,Z\ge y|X\le Z)}\cdot\frac{\Pr(X\le x,Z\ge y|X\le Z)}{\Pr(X\le x,Z=y,\delta=1|X\le Z)}$$
$$=\frac{\Pr(X=x,Y=y,C>y|X\le Z)}{\Pr(X=x,Y\ge y,C>y|X\le Z)}\cdot\frac{\Pr(X\le x,Y\ge y,C>y|X\le Z)}{\Pr(X\le x,Y=y,C>y|X\le Z)}$$
$$=\frac{\Pr(X=x,Y=y|X\le Y)}{\Pr(X=x,Y\ge y|X\le Y)}\cdot\frac{\Pr(X\le x,Y\ge y|X\le Y)}{\Pr(X\le x,Y=y|X\le Y)}$$
$$=1.\quad(\text{Under }H_0)$$

Under H_0 and Assumption B, all entries in Table 2 is observed under the conditioning event $X\le Y$ since $\Pr(X\le C)=1$ holds. Thus,

$$\frac{\Pr(X=x,Z=y,\delta=1|X\le Y)}{\Pr(X=x,Z\ge y|X\le Y)}\cdot\frac{\Pr(X\le x,Z\ge y|X\le Y)}{\Pr(X\le x,Z=y,\delta=1|X\le Y)}$$
$$=\frac{\Pr(X=x,Y=y|X\le Y)}{\Pr(X=x,Y\ge y|X\le Y)}\cdot\frac{\int_{u=0}^{x}\Pr(X=u,Y\ge y|X\le Y)\Pr(C_R>y-u)}{\int_{u=0}^{x}\Pr(X=u,Y=y|X\le Y)\Pr(C_R>y-u)}\quad\text{(B1)}$$
$$=\frac{\mathrm{d}F_X(x)\{-\mathrm{d}S_Y(y)\}}{\mathrm{d}F_X(x)S_Y(y)}\cdot\frac{\int_{u=0}^{x}\mathrm{d}F_X(u)S_Y(y)\Pr(C_R>y-u)}{\int_{u=0}^{x}\mathrm{d}F_X(u)\{-\mathrm{d}S_Y(y)\}\Pr(C_R>y-u)}\quad(\text{under }H_0)$$
$$=1.\tag{B.1}$$

Appendix C. Derivations of equivalent expressions

In this section, we prove Eqs. (6) and (17). Note that Eq. (6) is the uncensored case with $C_i=\infty$ in (17). For mathematical convenience, we define the discordant indicator $\Delta_{ij}=I\{(X_i-X_j)(Z_i-Z_j)<0\}$. To simplify the notations, let $W(\check{X}_{ij},\tilde{Z}_{ij})=\tilde{W}_{ij}$ and $R(\check{X}_{ij},\tilde{Z}_{ij})=\tilde{R}_{ij}$. One can write

$$L_w=\sum_i\sum_{\substack{j:X_j\le X_i\\X_i\le Z_j\le Z_i}}\delta_jW(X_i,Z_j)\left\{N_{11}(dX_i,dZ_j)-\frac{1}{R(X_i,Z_j)}\right\}$$
$$=\sum_i\delta_iW(X_i,Z_i)\frac{R(X_i,Z_i)-1}{R(X_i,Z_i)}-\sum_i\sum_{\substack{j:X_j<X_i\\X_i\le Z_j<Z_i}}\delta_jW(X_i,Z_j)\frac{1}{R(X_i,Z_j)}$$
$$\equiv I_1-I_2.$$

Using the fact that $\sum_j I(X_j<X_i,Z_j>Z_i)=R(X_i,Z_i)-1$, it follows that

$$I_1=\sum_i\sum_{j:X_j<X_i,Z_j>Z_i}\delta_i\frac{W(X_i,Z_i)}{R(X_i,Z_i)}=\sum_i\sum_{j:X_j<X_i,Z_i<Z_j}\delta_i\frac{\tilde{W}_{ij}}{\tilde{R}_{ij}}.$$

The indicator Δ_{ij} equals zero for a pair (i, j) with $X_j < X_i, Z_i < Z_j$. Therefore

$$I_1 = \sum_i \sum_{j:X_j<X_i,Z_i<Z_j} \delta_i \Delta_{ij} \frac{\tilde{W}_{ij}}{\tilde{R}_{ij}} = \sum_i \sum_{j:X_j<X_i,X_i<Z_j} \delta_i \Delta_{ij} \frac{\tilde{W}_{ij}}{\tilde{R}_{ij}}.$$

By applying similar algebraic manipulations, it follows that

$$I_2 = \sum_i \sum_{\substack{j:X_j<X_i \\ X_i\le Z_j<Z_i}} \delta_j \frac{\tilde{W}_{ij}}{\tilde{R}_{ij}} = \sum_i \sum_{j:X_j<X_i,X_i\le Z_j} \delta_j (1-\Delta_{ij}) \frac{\tilde{W}_{ij}}{\tilde{R}_{ij}}.$$

Combining I_1 and I_2, we obtain

$$L_W = \sum_i \sum_{j:X_j<X_i,X_i\le Z_j} \tilde{W}_{ij} \frac{\delta_i \Delta_{ij} - \delta_j(1-\Delta_{ij})}{\tilde{R}_{ij}} = \sum_i \sum_{j:X_j<X_i} I\{\check{X}_{ij} \le \tilde{Z}_{ij}\} \tilde{W}_{ij} \frac{\delta_i \Delta_{ij} - \delta_j(1-\Delta_{ij})}{\tilde{R}_{ij}}.$$

For a pair (i, j) with $X_j < X_i$, the following equation holds:

$$\delta_i \Delta_{ij} - \delta_j(1-\Delta_{ij}) = I\{(\delta_i = \delta_j = 1) \cup (Z_j - Z_i > 0 \,\&\, \delta_i = 1 \,\&\, \delta_j = 0) \cup (Z_i - Z_j > 0 \,\&\, \delta_i = 0 \,\&\, \delta_j = 1)\}(2\Delta_{ij} - 1).$$

Thus, we obtain Eq. (17) as follows:

$$L_W = \sum_i \sum_{j:X_j<X_i} I\{B_{ij}\} \tilde{W}_{ij} \frac{2\Delta_{ij}-1}{\tilde{R}_{ij}} = \sum_{i<j} I\{B_{ij}\} \tilde{W}_{ij} \frac{2\Delta_{ij}-1}{\tilde{R}_{ij}}$$

$$= -\sum_{i<j} I\{B_{ij}\} \frac{\tilde{W}_{ij}}{\tilde{R}_{ij}} \mathrm{sgn}\{(X_i - X_j)(Z_i - Z_j)\}.$$

The second equation follows from the permutation symmetry of each term with respect to arguments (i, j).

References

[1] M.C. Wang, Nonparametric estimation from cross-sectional survival data, Journal of the American Statistical Association 86 (1991) 130–143.
[2] S.W. Lagakos, L.M. Barraj, V. De Gruttola, Non-parametric analysis of truncated survival data, with application to AIDS, Biometrika 75 (1988) 515–523.
[3] J. Hyde, Testing survival under right censoring and left truncation, Biometrika 64 (1977) 225–230.
[4] J. Hyde, Survival analysis with incomplete observations, in: R.G. Miller, B. Efron, B.W. Brown, L.E. Moses (Eds.), Biostatistics Casebook, John Wiley and Son, New York, 1980, pp. 31–46.
[5] J.P. Klein, M.L. Moeschberger, Survival Analysis: Techniques for Censored and Truncated Data, Springer, New York, 2003.
[6] D. Lynden-Bell, A method of allowing for known observational selection in small samples applied to 3rc quasars, Mon. Nat. R. Astr. Soc. 155 (1971) 95–118.
[7] M.C. Wang, N.P. Jewell, W.Y. Tsai, Asymptotic properties of the product-limit estimate and right censored data, Annals of Statistics 13 (1986) 1597–1605.
[8] M. Woodroofe, Estimating a distribution function with truncated data, The Annals of Statistics 13 (1985) 163–177.
[9] W.-Y. Tsai, Testing the assumption of independence of truncation time and failure time, Biometrika 77 (1990) 169–177.
[10] L. Chaieb, L.-P. Rivest, B. Abdous, Estimating survival under a dependent truncation, Biometrika 93 (2006) 655–669.
[11] E.C. Martin, R.A. Betensky, Testing quasi-independence of failure and truncation via conditional Kendall's tau, Journal of the American Statistical Association 100 (2005) 484–492.
[12] C.-H. Chen, W.-Y. Tsai, W.-H. Chao, The product-moment correlation coefficient and linear regression for truncated data, Journal of the American Statistical Association 91 (1996) 1181–1186.
[13] T.R. Fleming, D.P. Harrington, Counting Process and Survival Analysis, John Wiley and Sons, New York, 1991.
[14] D.P. Harrington, T.R. Fleming, A class of rank test procedures for censored survival data, Biometrika 69 (1982) 133–143.
[15] D.G. Clayton, A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, Biometrika 65 (1978) 141–151.
[16] D. Oakes, Semi-parametric inference in a model for association in bivariate survival data, Biometrika 73 (1986) 353–361.
[17] D. Oakes, Bivariate survival models induced by frailties, Journal of the American Statistical Association 84 (1989) 487–493.
[18] C. Genest, R.J. Mackay, The joy of Copulas: Bivariate distributions with uniform marginals, The American Statistician 40 (1986) 280–283.
[19] S. He, G.L. Yang, Estimation of the truncation probability in the random truncation model, Annals of Statistics 26 (1998) 1011–1027.
[20] C. Genest, Frank's family of bivariate distributions, Biometrika 74 (1987) 549–555.
[21] T. Emura, W. Wang, Asymptotic Analysis and Variance Estimation for Testing Quasi-independence under Truncation, Technical Reports of Mathematical Sciences, Chiba University, vol. 24, No. 12. 2008. Available at http://www.math.s.chiba-u.ac.jp/report/files/08012.pdf.
[22] J. Shao, Differentiability of statistical functionals and consistency of the jackknife, The Annals of Statistics 21 (1993) 61–75.
[23] W.Y. Tsai, N.P. Jewell, M.C. Wang, A note on the product-limit estimator under right censoring and left truncation, Biometrika 74 (1987) 883–886.
[24] J. Cuzick, Rank tests for association with right censored data, Biometrika 69 (1982) 351–364.
[25] J. Cuzick, Asymptotic properties of censored linear rank tests, The Annals of Statistics 13 (1985) 133–141.
[26] D.M. Dabrowska, Rank tests for independence for bivariate censored data, The Annals of Statistics 14 (1986) 250–264.
[27] D. Oakes, A concordance test for independence in the presence of censoring, Biometrics 38 (1982) 451–455.
[28] J.H. Shih, T.A. Louis, Tests of independence for bivariate survival data, Biometrics 52 (1996) 1440–1449.
[29] J.H. Shih, T.A. Louis, Inference on the association parameter in copula models for bivariate survival data, Biometrics 51 (1995) 1384–1399.
[30] L. Hsu, R.L. Prentice, A generalisation of the Mantel–Haenszel test to bivariate failure time data, Biometrika 83 (1996) 905–911.
[31] A.A. Ding, W. Wang, Testing independence for bivariate current status data, Journal of the American Statistical Association 99 (2004) 145–155.
[32] D. Beaudoin, L. Lakhal-Chaieb, Archimedean Copula model selection under dependent truncation, Statistics in Medicine 27 (2008) 4440–4454.
[33] A.W. Van Der Vaart, Asymptotic Statistics, in: Cambridge Series in Statistics and Probabilistic Mathematics, Cambridge University Press, Cambridge, 1998.